structural indicators of the extent of competition that are motivated by the S/C/P industrial organization paradigm.

Three phases are identified, namely, the phases of (1) contested industry reorganization; (2) market expansion and growth; and (3) market maturity. The role of regulation is indicated for each phase.[88]

Contested Industry Reorganization. The phase of contested industry reorganization begins as a consequence of a "paradigm shift" that results in the creation of an entirely new industry or the establishment of a concept or policy viewpoint, such as competition, that threatens the existing organization of industry. Decisions by the FCC permitting entry of "specialized common carriers" into the interstate private line market and the connection of "foreign attachments" to the public switched network represented such paradigm shifts favoring competition, respectively, in the long distance and customer premises equipment markets. During this phase, incumbent dominant firms challenge the legitimacy of the paradigm shift and adopt an adversarial, even hostile, posture toward firms, government, or other groups that advocate a new view of industry organization.

New property rights emerge during this phase that represent direct implications of the new paradigm. The establishment of these new property rights requires extensive regulatory rulemaking and often protracted judicial review. Regulatory activity is intense, resulting in new regulatory mechanisms that were unnecessary prior to the paradigm shift favoring competition. Dominant firm strategic behavior is focused on deterring or discouraging the entry of rivals. Eventually, a landmark regulatory ruling or court decision, as previously discussed, makes definitive the property rights that are essential to reducing the risk of investment by entrants.

At present, competition in markets for local telecommunications services falls within this stage. Property rights in enhanced interconnection with LECs and central office collocation are not definitively resolved. In some states, direct competition with LECs is unlawful. Although some LECs wish to merge with other telecommunications firms, the underlying economic viability of such integration strategies is uncertain.[89] Until definitive decisions are reached in law or regulation concerning property rights, deregulation of LECs during this phase will result in dominant carrier strategic conduct that is inimical to the eventual development of

---

[88] The notion of phases is an adaptation of Shepherd's "stages" in the life-cycle of firms defined as public utilities. See William G. Shepherd, *The Treatment of Market Power* (New York: Columbia University Press, 1975), Chapter 9.

[89] See the discussion of this point in Section 3.

competition. Traditional regulatory mechanisms, such as price cap regulation, remain essential for constraining certain types of dominant firm strategic pricing behavior that might otherwise deter competitive entry or disadvantage rivals.

Market Expansion and Growth. The phase of market expansion and growth emerges after uncertainty surrounding crucial property rights is removed. As previously discussed, pivotal decisions by the FCC that were upheld on judicial review made competition possible in both the terminal equipment and long distance telecommunications markets. Similarly, the AT&T divestiture also represented a defining event in the development of telecommunications competition that made market entry more attractive to rivals.

During this phase, the entry of new firms stimulate the growth of the entire industry, such that the revenues of the dominant carrier may still continue to grow, although the market share of the dominant carrier begins to decline. So long as regulatory authorities remain sensitive to possible dominant firm strategic conduct, some modification of the regulatory paradigm applied to dominant carriers may be possible. Regulatory processes for addressing possible transactional barriers should be in place before streamlined regulation of dominant carriers is fully implemented.

Market Maturity. Although market share rules are a problematic measure of the extent of competition, they provide one possible indicator of the diminution of dominant carrier market power. As the market share of a dominant carrier approaches 40-50 percent, the potential for effective strategic behavior that might threaten rivals is limited. Substantial deregulation of dominant carriers, i.e., the discontinuance of price cap regulation, is appropriate during this phase. Some regulations should continue indefinitely, such as rules regulating network interfaces, interconnection, and access to local exchange facilities. Such regulations, like the Uniform Commercial Code that governs ordinary market exchange in otherwise "unregulated markets," provides the essential, ongoing legal infrastructure that is key to designing and implementing efficient transactional modes in the telecommunications industry.

3 Outlooks for Industry Evolution

The discussion in the previous section of this paper has provided a historical and conceptual analysis that establishes the basis for both identifying and treating potential transactional barriers to competition in local telecommunications markets. These barriers are much more troublesome to the extent that the LECs continue to produce services and possess facilities that are unique, because competitors to the LECs may depend upon these resources for the creation of competitive offerings.

This section of the paper surveys the diverse array of emerging technologies that might permit substitution for the local telecommunications services offered by the LECs. The evaluation indicates that each of the major types of networks may be superior for certain classes of traffic. Hence, substitution of these other technologies for LEC networks and facilities appears unlikely, and dependence upon the LECs for certain essential features and functions may continue.

This analysis indicates that a very complex industry structure may evolve in local telecommunications markets. The possible diversity of local technologies and services, some of which may be interdependent, underscores the critical importance of understanding the underlying factors that will influence transaction costs—especially asset specificity—in future exchange relationships between the LECs and their rivals. Absent such an understanding, and appropriate regulatory policy and processes, an inefficient industry structure may emerge as a consequence of unaddressed transactional barriers.

3.1 A Comparison of Industry Structures for Interexchange and Local Telecommunications Services

If, using only their own facilities, different competing carriers can produce the same service for the same cost, then none of the carriers possess essential facilities needed by the others. This is basically the case for the long-haul component of interexchange competition. The essential ingredient needed by all interexchange carriers—equal access—was obtained under the terms of the decree that divested the Bell Operating Companies from AT&T. Coincidental with this legal and policy event, two technological developments had evolved to the point of commercial viability: fiber optic transmission and low cost digital switching. The major facilities-based interexchange carriers quickly took advantage of these technologies. Rights-of-ways were also required, but these were obtained along railroad lines and highways.

As a result of the conjunction of these regulatory and technological changes, there now exists three national facilities-based carriers, other regional carriers, and other more specialized carriers, all using similar technologies and all producing similar services. Because of the increasing "commodity like" nature of interexchange service, carriers now attempt to distinguish themselves from one another by value-added features, such as billing, that reflect an industry sector directly responsive to market pressures.

Unlike the interexchange case, the services produced by different types of local telecommunications networks are likely to remain very diverse and not substitutable. This is the conclusion of the following analysis, which examines cable television networks as a substitute for LEC local networks and radio-based networks as a substitute for all landline-based networks.

If different carriers can only produce different services, or the same services at substantially different costs, then certain of these carriers may possess essential facilities. If these facilities were obtained under a monopoly grant and if policy is now changing to favor competition, then allowing other carriers access to these facilities may be necessary for such competition to be effective. The situation just described applies to local exchange competition, with the LECs possessing essential facilities originally obtained as a monopoly award.

Assertions that LEC networks and cable television networks are "converging" are now fashionable. In reality, these two types of networks may evolve to a common ground very slowly, with each continuing to reflect its individual heritage for a long time. Another commonly-held view is that wireless communications—based upon traditional cellular systems or more recent PCS technology—will soon substitute for those services offered by landline networks. This projection is on a collision course with another popular contention: that high-datarate services such as interactive multimedia and video on demand will prevail in the future. The radio spectrum that would be required for high-datarate services may relegate the majority of these offerings to landline networks. Both of these matters are now discussed.

3.2 The Likely Separation of Cable TV and LEC Networks

Traditional cable television networks use a so-called "tree and branch" architecture, which is optimized to deliver signals on a point-to-multipoint or broadcast basis. In this design, a single signal originates from a central point and then radiates out multiple branches—very much in the image of a tree with a trunk, branches and then individual leaves.

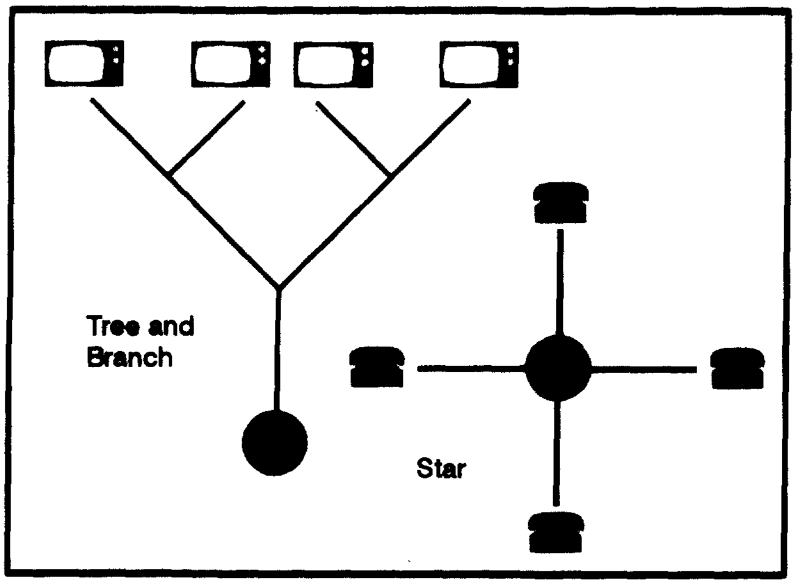

Figure 5—Comparison of Tree-and-Branch and Star Networks

Traditional LEC networks, on the other hand, are organized as "star" networks, with individual circuits emanating from a central office building and terminating, individually, at the locations of served customers. This architecture is structured for point-to-point traffic. In contrast to cable television networks, LEC networks also involve switching at the central office. This switching can connect any two of the circuits served by the central office or can connect any one of the served circuits to a distant central office building. Figure 5 contrasts these two architectures.

Because of the signal attenuation characteristics of coaxial cable, traditional cable television systems may require several amplifiers for every mile of plant. A long cascade of such amplifiers may thus be interposed between the point of signal origination and the customer. Such cascades cause a number of problems and limitations.

To resolve these technical difficulties, cable television systems are beginning to install optical fiber in the trunk portions of their plants. Traditional coaxial cable is still used in the branches, but the number of cascaded amplifiers are reduced. Such hybrid systems can improve signal quality, reduce maintenance costs and increase the number of traditional channels that can be supplied to individual customers. While such systems may move in the direction of emerging LEC local distribution architectures, these hybrid cable systems still deliver only point-to-multipoint traffic and still do not involve switching.[90]

While the vast majority of fiber-optic investment in the cable industry is in networks such as those described above, some experimentation and near-commercial deployment is also taking place with more advanced systems. For example, Time Warner plans to install a system that will support point-to-point, broadband, switched services near Orlando, Florida. This system, which initially is said to be planned for 4,000 customers, reportedly can provide ". . . video on demand, interactive games, distance learning, full-motion video interactive shopping and access to interexchange carrier networks."[91]

The various LECs do not appear consistent in their approaches toward the introduction of optical fiber into their local distribution plants. Pacific Bell has announced very aggressive plans to introduce a hybrid fiber optic and coaxial cable broadband network which it claims will provide voice, data and video-on-demand services on an integrated basis to 5.5 million customers by the year 2000.[92] Bell

---

[90] For a discussion of the current use of optical fiber in cable television plants and possible future architectures, see Gary Kim, "Cable TV Facilities as Integrated Broadband Networks," *Fiber Optics Magazine* (March 1991):16-26.

[91] "Time Warner plans switched, broadband cable TV system in Florida, will compete with BellSouth; IXC access, PCS, video-on-demand, interactive video, other services planned," *Telecommunications Reports* 59 (February 1, 1993):37-38.

[92] "Pacific Bell unveils plan to deploy broadband local distribution network to 5.5 million

Atlantic, on the other hand, seems more tentative and is experimenting with several different strategies. In Virginia, they are examining Asymmetric Digital Subscriber Line (ADSL) technology, which can provide a compressed video-on-demand signal over traditional copper local loop plant.[93] In New Jersey, however, they reportedly have ordered a "fiber to the curb" systems adequate to serve 100,000 access lines.[94] These systems are intended to provide both video and switched voice services. Ameritech apparently plans to retain its current metallic loop plant for voice services, but to overlay this network with a hybrid optical fiber and coaxial cable system that will supply video programming and interactive services. Ameritech reportedly plans to deploy this system at the rate of one million lines per year for fifteen years.[95]

In summary, the cable television industry is introducing optical fiber into its plant, but mainly on the justification of reduced cost and an improvement in traditional services. Among the LECs, Pacific Bell has announced plans for an integrated network, Ameritech plans separate networks for traditional voice and video, and Bell Atlantic's plans regarding network integration appear uncertain.

Johnson and Reed have performed a theoretical analysis that questions the economic feasibility of integrated broadband networks.[96] They postulate a modern hybrid fiber/metallic loop "LEC type" local distribution plant to supply switched voice, and a separate modern hybrid fiber/coaxial "CATV type" plant to supply traditional cable television services. Both systems are assumed to be installed in a large new residential community. The sum of the investment costs for these two networks, per home passed, is \$568 + \$368 = \$936. The authors contrast this with an integrated system that can supply both narrowband and broadband services on a point-to-point switched basis. This system has an initial investment, per home passed, of \$1,760. They question whether the additional investment required for the integrated system (\$1,760 - \$936 = \$824) is justified in relation to the only additional

---

homes by year 2000," *Telecommunications Reports* 59 (November 15, 1993):1-3.

[93] "Bell Atlantic plans video-on-demand trial using 'ADSL' technology to deliver video over copper lines; C&P of Virginia files section 214/'video dial-tone' application," *Telecommunications Reports* 58 (October 26, 1992):10-11.

[94] "Bell Atlantic orders 'fiber-to-the-curb' systems from BBT," *Telecommunications Reports* 59 (May 3, 1993):20-21.

[95] "Ameritech Unveils Broadband Video Platform Deployment Plan," *Telecommunications Reports* 60 (January 31, 1994):9.

[96] Leland L. Johnson and David P. Reed, "Telephone Company Entry into Cable Television," *Telecommunications Policy* (March 1992):122-134.

service that they believe such an integrated system can provide, which is video on demand.

As a separate exercise, Johnson and Reed then remove the video server and video switching from their hypothetical integrated system, creating an integrated local distribution network that can still supply traditional switched voice and traditional cable television services. The cost of this simplified integrated system is $1,449 per home passed. This is significantly higher than $936, the sum of the costs for the two separate networks that together can supply the same two services. The authors conclude that

> . . . [T]he tree and branch architecture used today by cable operators is an efficient mechanism for distributing video services. In contrast, efficient delivery of switched services calls for some variant of a star network. Combining the two services onto a single system can actually add to costs. In other words, economies of scope seem not to exist in bringing together broadband and narrowband residential services. Contrary to the implications of some descriptions of fiber to the home, there is nothing magical about service 'integration'.[97]

Concerning the possibility of LEC networks and cable television networks converging, the actual industry experience discussed above presents a very mixed picture, while the Johnson and Reed model argues that such convergence would not be efficient. None of this, of course, proves that convergence will not occur. But the combined weight of all this evidence does suggest that, in many geographic markets, LEC networks and cable television networks may not be combined.

This last observation points to another way in which local exchange competition is more complex than has been the case for competition in customer premises equipment or interexchange services. There is a *geographical* dimension to local telecommunications competition that is absent in the other two cases. Customer premises equipment is sold in a national market, as are interexchange services (at least for the three major facilities-based carriers). Identical equipment and services are sold in New York city as in Possum Gap, West Virginia. For local exchange competition, in contrast, different market structures may evolve in different places.

3.3 The Possible Divergence of Landline and Radio Based Services

High-datarate services, such as interactive multimedia and video on demand, are often discussed as future possibilities for both residential and business applications. The promise of offering such services is a primary reason for LECs evolving toward fiber-based, point-to-point, switched, broadband networks. These services are

[97] *Ibid.*, p. 129.

technically feasible on such landline networks, questions of economic feasibility aside. Another major direction in future telecommunications services is toward the use of radio links instead of landline links. Radio-based, "untethered" communication can increase both user functionality and convenience. There may be a major conflict between this trend toward radio and the trend toward higher datarate services. The objective of providing high datarate services on radio-based networks raises questions about spectrum availability, since much more spectrum would be required than for traditional voice offerings. If adequate spectrum is not available, then high datarate services may be restricted mainly to landline networks.

Interactive multimedia presentations may involve audio, text, computer graphics, realistic still pictures and motion video. The user interacts with such material, guiding the presentation alone a unique path. A multimedia show does not necessarily require motion video. For example, the current implementation of the Prodigy on-line system—which provides text, computer graphics and realistic still images—would qualify as a multimedia offering under this definition. But most future multimedia systems of real interest are assumed to include motion video as an element. While the datarate from the central serving node to the user must be high enough to support such video images, the datarate in the opposite direction, from the user to the node, may be quite low. Only a modest datarate may be necessary to convey the information inherent in the actions of user selection.

Video on demand has similar datarate requirements and exhibits a similar asymmetry. In this service, a user can call up a traditional video presentation (e.g., a movie) from a central serving node at an arbitrarily chosen time. Unlike interactive multimedia, the presentation is linear and predetermined. The datarate from the node to the user would be similar to that for multimedia (i.e., adequate to support video); the datarate from the user back to the node might be even less than multimedia would require.

Consider a small, radio-linked device such as a portable interactive multimedia terminal or video-on-demand receiver. Such equipment requires a unique, point-to-point video signal. A datarate of 1.5 million bits per second (mbps) might be adequate for this application.[98] Such traffic primarily is one way, from the central

[98] An uncompressed standard television (NTSC) signal requires a datarate of about 100 mbps. Extensive work by the Motion Picture Expert Group (MPEG) has produced a standard for compressing an NTSC signal to about 1.5 mbps. A signal compressed to this degree is visually inferior to an uncompressed signal, but might be adequate for a small-screen, portable device like the one discussed here. (The datarate of 1.5 mbps is consistent with that available from first-generation CD-ROM drives, which could be used to store such compressed video information. The datarate also is compatible with

serving node to the portable receiver, with only a low datarate control channel in the opposite direction. In contrast, a portable telephone is assumed to generate two-way traffic at 16 thousand bits per second (kbps) in each direction.[99] Dividing these two numbers yields: 1.5 mbps/16 kbps = 93.75. Compensating for the two-way nature of the portable telephone produces: 93.75/2 = 46.875, or about 50 as the ratio of the gross datarates required for video versus voice.

Thus, using digital coding, about fifty times as much spectrum may be required to support a portable multimedia terminal or video-on-demand receiver than is required to support a standard portable voice telephone. This spectrum demand may place severe limitations on such portable video applications.

Research for the third generation of mobile communication systems is now underway (counting analog cellular systems as the first generation and the still unlicensed and undeployed digital PCS systems as the second). If the popularity of mobile communications continues to grow as projected, the traffic demand will cause third-generation systems to employ a coordinated overlay of traditional macrocells, smaller microcells and tiny, indoor picocells with diameters down to 100 meters. Such arrangements will permit the level of frequency reuse required to handle the traffic within the spectrum that may be available. At high user demand, this organization would be necessary even for all-voice traffic. These configurations will require the development of very inexpensive base stations and will generate a high volume of handoffs with an attendant burden on the switching and control infrastructure.[100]

Third generation systems will represent the first attempt to integrate mobile voice, data and video. Higher datarate services such as video may only be available ". . . either within a special environment (e.g., within a business customer premises

---

first-generation Asymmetric Digital Subscriber Line technology, which could be used to transmit such a compressed video signal over standard LEC loop plant.) See T. Russell Hsing, Cheng-Tie Chen and Jules A. Bellisio, "Video Communications and Services in the Copper Loop," *IEEE Communications Magazine* 31 (January 1993):62-68.

[99] Possibilities for digital speech encoding over radio channels range from high-datarate schemes (such as adaptive pulse code modulation at 32 kbps) down to relatively low datarate methods at 8 kbps or below. Because low datarate approaches are more susceptible to errors during transmission, error detection and correction means must be used in such cases. Balancing these factors, an assumption of 16 kbps seems reasonable. For a discussion of the need for higher datarates to maintain voice quality, see Filip Lindell, Johan Sköld, Per Willars and Erik Nilsson, "Radio Access Technology Evolution," *Ericsson Review* 70 (n3, 1993):85.

[100] David J. Goodman, "Cellular Packet Communications," *IEEE Transactions on Communications* 38 (August 1990):1274.

network), or on the basis of bandwidth reservation in a public network."[101] In a business or professional environment, tiny picocells might be employed within the confines of a single floor of a building. This arrangement would permit a high level of frequency reuse, allowing adequate spectrum to be dedicated to an individual video signal. Bandwidth reservation on a public network would require the dedication of considerable common-carrier resources for a predetermined interval for the support of an individual portable video terminal. Cost considerations might restrict both such business and public network applications to special-purpose situations. Thus, for most of the traffic, higher datarate services may be restricted to landline networks. If so, then landline-based and radio-based local networks may provide substantially different types of services in the future.

4 Conclusions

The analysis of this paper supports specific conclusions concerning (1) economic methodology for studying the extent of competition in telecommunications markets; (2) the establishment of property rights and the emergence of competition in the U.S. telecommunications industry; (3) the possible effects of transaction cost considerations on the performance of local telecommunications services markets; and (4) the appropriate role of federal regulation during the early phases of the development of competition in markets for local telecommunications services.

Concerning economic methodology for studying the extent of competition in telecommunications markets, the paper supports the following conclusion:

1. Application of economic concepts developed within the literature of the New Institutional Economics, specifically, the economics of property rights and transaction cost economics, complements the structure-conduct-performance paradigm as a method for assessing the extent of competition by explaining how transactional barriers may affect voluntary market exchange.

Concerning the relationship of property rights and the development of competitive telecommunications markets, the paper supports the following conclusions:

2. The establishment of property rights in interconnection with and access to incumbent LECs is a necessary condition (although not necessarily sufficient) for the development of competition in U.S. telecommunications markets.

3. Although property rights have important implications for achieving an economically-efficient allocation of resources and fostering the emergence of competitive markets, the process of identifying, defining, and implementing such rights involves extensive legal and regulatory proceedings. Since such processes are constrained by law

[101] Stanley Chia, "The Universal Mobile Telecommunication System," *IEEE Communications Magazine* 30 (December 1992):55.

and the Constitution to provide due process, the establishment of new property rights to permit competition to develop necessarily takes a long time, i.e., years, not months.

4. Property rights affecting economic transactions in particular and market exchange in general change through time. Industry structure and organization based on an existing property rights paradigm will change as a consequence of the emergence of new property rights.

5. During the early years of the emergence of competition in telecommunications markets, regulatory commissions should be alert to ill-defined property rights and initiate appropriate proceedings to identify, define, and enforce such rights as may be necessary to establish these critical prerequisites for competitive markets.

Concerning the possible effects of transaction cost considerations on the performance of local telecommunications services markets, the paper supports the following conclusions:

6. Strategic behavior by LECs may impede market exchange with rivals. Examples of such behavior may include defining an LEC service or functionality requested by a competitor in such a way to disadvantage the rival in addition to other types of dominant firm conduct that raises rivals' cost of production. There exist circumstances, however, where impediments to market exchange between LECs and their rivals may be attributable to unique attributes of the transaction and not LEC strategic behavior.

7. Certain attributes of a transaction between trading parties, e.g., a LEC and its rivals, will affect the efficacy of market exchange, resulting in the integration of firms if transactional barriers are sufficiently high. Asset specificity is the key attribute of a transaction that will affect the efficacy of market exchange, although both uncertainty and transaction frequency play important roles.

Concerning the appropriate role of federal regulation during the early phases of the development of competition in markets for local telecommunications services, the paper supports the following conclusions:

8. The choice of general purpose versus special purpose technology by LECs may represent a strategic decision to deepen asset specificity that, in turn, may provide a rationale for integration. Regulatory scrutiny of LEC investment decisions is essential to determine what effect the choice of technology will have on transaction cost and the sustainability of market exchange and competition.

9. Even if the emerging pattern of consumer demand for LEC services and facilities requires more LEC investment in transaction-specific assets, market exchange may still be the most efficient transactional mode *if* regulatory authorities establish both rules and processes, formal and informal, for resolving *ex post* transactional disputes. Such regulatory rules and processes are viewed in a very different way than is customary: such regulatory activity *reduces the transaction cost of using markets* in the face of possible contracting problems implied by deepening asset specificity. As a result, the powerful incentives provided by a rivalrous market to reduce the cost of production and provide consumers with the goods and services that are wanted are preserved.

10. As the local telecommunications service industry evolves, the FCC may need to define new property rights that are currently undefined or latent within the public utility

concept. As government policy evolves toward an open entry industry model, rights and obligations inherent in the former industry structure must evolve as well.

11. Broadly speaking, competition in local telecommunications service markets is in a phase of contested industry reorganization, where property rights essential for competitive markets remain incompletely defined. Unlike the development of competition in the nationwide market for long distance telecommunications, firms supplying the market for local telecommunications services will utilize different technologies; offer differentiated services that are both complementary with and substitutable for those currently provided by LECs; and will require more complex and differentiated types of local exchange interconnection compared to that required by long distance carriers. Intensive involvement of the FCC in the development of appropriate pro-competitive regulatory policy, not deregulation of LECs or streamlined price cap rules, is required to define necessary property rights and to minimize transactional barriers during this phase of the development of competition in local telecommunications services markets.

Jerry B. Duvall is a consulting economist in private practice in Germantown, Maryland. He holds a Ph.D. in Economics and has previously been employed by Intelsat and the Federal Communications Commission.

John G. Williams is a Vice President of the Telecommunications Consulting Group, Inc. He holds a Ph.D. in Computer Science and has previously been employed by the Federal Communications Commission and Bell Laboratories.

Appendix A
The Structure-Conduct-Performance Paradigm

This Appendix provides a more detailed description of the *structure-conduct-performance* (S/C/P) paradigm as introduced and then briefly discussed in Section 2.1 of the paper.[102] Figure 2 in Section 2.1 shows that certain *basic conditions* determine *market structure*. Market structure, in turn, influences the *conduct* of business firms. The interaction of market structure and conduct, represented both by the solid line segments that denote direct causal links, and by the broken line segments that denote "feedback" effects, determine the level of *market performance* observed in the market.

Basic conditions include factors that influence both the supply and demand side of the market. For example, the location and ownership of essential raw materials; the nature of available technology; the degree of labor force unionization; the durability of the product; business attitudes; and other such basic conditions will affect the supply side of a market. The extent of price elasticity of demand; the availability of substitute products or services; the rate of growth in product demand; marketing characteristics of the product or service; and other such basic conditions will influence the demand side of the market. The nature of laws and government policies affecting a particular market are other basic conditions that may be expected to affect both the supply and demand side of the market.

Market structure includes such elements as the number of buyers and sellers interacting in the market; the nature and extent of product differentiation; barriers to entry and exit from the market; the capital intensity of production that affects the ratio of fixed and variable costs; and the extent of vertical integration of firms supplying the market. *Conduct* refers to the type of pricing behavior observed in the market; the nature of product strategy and advertising undertaken by firms supplying the market; the nature and extent of research and innovation pursued by firms in the industry; the extent of plant investment; and legal tactics employed by firms.

Performance refers to an assessment of how well a given market or industry seems to function. The specific evaluative criteria used in making such an assessment usually include static and dynamic economic efficiency; equity; and, sometimes, macroeconomic criteria, such as the achievement of full employment. The notions of

[102] The discussion of this Appendix closely follows F. M. Scherer and David Ross, *Industrial Market Structure and Economic Performance,* 3rd ed. (Boston: Houghton Mifflin, 1990), Chapter 1.

static and dynamic economic efficiency are usually stressed by economists in assessing the performance of any given market. Static economic efficiency as a criterion of market performance is achieved if it is impossible to improve the economic well-being of any market participant (buyer or seller) through a reallocation of resources within the industry without simultaneously adversely affecting the economic well-being of some other market participant.[103] If the economic well-being of one market participant can be improved without harming any other buyer or seller, the resulting gain in static economic efficiency is referred to as a *Pareto improvement* in economic welfare. Static economic efficiency in the production of goods and services within the industry and the consumption of the output of the industry by consumers is fully realized if all opportunities for achieving Pareto improvements are exhausted.

The notion of static economic efficiency in either production or consumption implicitly takes the current state of technology as a given. Thus, static economic efficiency as an indicator of market performance is defined with reference to a given state of technology at a specific time. Over time, technical knowledge accumulates and improves, making possible both "product innovations," i.e., new and improved products and services, and "process innovations," i.e., new and improved ways of producing goods and services, that effectively reduce the cost of production for any level of output. Innovations represent the successful *application* of changes and improvements in technical knowledge.

Dynamic economic efficiency as an indicator of market performance refers to the extent to which technical change is absorbed and implemented through product and process innovations by firms supplying the market.[104] The realization of dynamic economic efficiency means that firms not only minimize the cost of production *today* but also pursue a strategy of adopting innovations that will lower the cost of production *through time* or result in the production of new or improved products and services. The criterion of dynamic economic efficiency as an indicator of good market performance is not met solely by implementing "new" technology per se: adoption of new technology that *raises* the cost of production is dynamically *inefficient*, unless it produces new benefits or incremental improvements in benefits to consumers that

---

[103] In more technical terms, static economic efficiency is referred to as *Pareto Optimality* and is a central concept in modern welfare economics. For further discussion, see Richard E. Just, Darrell L. Hueth, and Andrew Schmitz, *Applied Welfare Economics and Public Policy* (Englewood Cliffs, N.J.: Prentice-Hall, 1982), Chapter 2.

[104] A good discussion of dynamic economic efficiency is provided by Burton H. Klein, *Dynamic Economics* (Cambridge, Mass.: Harvard University Press, 1977), Chapter 3.

are at least *worth* the increased cost of production. Adoption of new technologies of production simply because they are available without forecasting their long-term effects on the incremental cost of production and the value of the incremental benefits to consumers is a strategy that is likely to conflict with the criterion of dynamic economic efficiency.

The concept of equity as an indicator of market performance can be defined in many ways and within the context of different disciplines, such as law, ethics, religion, economics, business administration, and others.[105] In most cases, equity as an attribute of market performance in private-sector markets refers to issues of fairness in the distribution of income between business firms and their customers. Thus, a business firm that faces little or no competition may be able to establish product or service prices well above the average cost of production, resulting in substantial "excess profits" or "economic rents" that accrue to the firm. Such an outcome may be considered as inequitable or unfair to the consumer.

In general, the S/C/P paradigm predicts that a *competitive market structure*, i.e., a market where "many" firms compete with each other for the business of "many" customers, will result in *competitive behavior* where no single firm has much influence, i.e., market power, over the prevailing market price or the total level of output produced and exchanged in the market. Rivalry between and among the many firms supplying the market tends to drive market price for the essentially homogeneous output toward the marginal cost of production, which eliminates "excess profits" that firms might otherwise realize in the absence of competitive rivalry. Competitive rivalry produces good market performance by providing consumers with the goods and services they prefer at unit prices that reflect the marginal cost of production.

The S/C/P paradigm further predicts that elements of market structure that restrict the number of firms supplying the market from "many" to "few" will diminish the extent of competitive rivalry among firms and foster a degree of *interdependence* of behavior among such firms. Such interdependence will dampen competitive pressures forcing market price toward the marginal cost of production and will adversely affect market performance in terms of the realization of static economic efficiency in both production and consumption. Often the relative

---

[105] The notion of equity within the context of regulated industries is especially complex. An illuminating discussion of equity viewed from an economic perspective is provided by Edward E. Zajac, "Perceived Economic Justice: The Example of Public Utility Regulation" in *Cost Allocation: Methods, Principles, Applications*, ed. H. Peyton Young (Amsterdam: North-Holland, 1985), pp. 119-153.

"fewness" of firms supplying this market is attributable to certain barriers to entry or attributes of the technology of production that effectively preclude the participation of many firms in the industry. Thus, a non-competitive industry structure results in non-competitive behavior by firms that, in turn, adversely affects market performance.

Transaction cost economics does not dispute the broad validity of the S/C/P paradigm as a methodology for assessing the extent of competition prevailing in a given market or industry. Rather, transaction cost economics provides an additional dimension to the study of competition by examining attributes of business transactions that may impede the development of competitive markets. To the extent that the S/C/P paradigm fails to recognize potential transactional barriers, an assessment of the extent of competition and market performance based solely on the S/C/P paradigm may be incomplete, or more seriously, possibly misleading in its implications concerning appropriate public policy.

Appendix B

Transaction Cost Economics: Further Discussion of Basic Concepts

This Appendix briefly discusses three basic ideas of transaction cost economics that supplement and extend the discussion of the text itself, namely, (1) *The Fundamental Transformation*; (2) *specific and nonspecific costs*; and (3) *Williamson's hostage concept*.

The Fundamental Transformation. The discussion in the text itself explicates the importance of asset specificity, uncertainty, and transaction frequency in making the transaction cost-minimizing choice between hierarchy or market exchange for effectuating transactions. This discussion did not emphasize, however, the potential change in the contracting environment after the contract is negotiated. So important are the implications of this potential environmental change that Williamson refers to it as *The Fundamental Transformation*.[106] Williamson sharply distinguishes between the *ex ante* contracting environment, i.e., the contracting environment before a transaction is completed, and the *ex post* contracting environment, i.e., the contracting environment after a contract has been negotiated and ratified by all parties. Williamson explains that

[106] Oliver E. Williamson, *The Economic Institutions of Capitalism* (New York: The Free Press, 1985), p. 61.

> Economists of all persuasions recognize that the terms upon which an initial bargain will be struck depend on whether noncollusive bids can be elicited from more than one qualified supplier. Monopolistic terms will obtain if there is only a single highly qualified supplier, while competitive terms will result if there are many. Transaction cost economics fully accepts this description of *ex ante* bidding competition but insists that the study of contracting be extended to include *ex post* features. Thus initial bidding merely sets the contracting process in motion. A full assessment requires that both contract execution and *ex post* competition at the contract renewal interval come under scrutiny.[107]

After the contract is signed, the contracting environment is fundamentally transformed. Again, Williamson explains that

> . . . [T]ransaction cost economics holds that a condition of large numbers bidding at the outset does not necessarily imply that a large numbers bidding condition will prevail thereafter. Whether *ex post* competition is fully efficacious or not depends on whether the good or service in question is supported by durable investments in transaction-specific human or physical assets. Where no such specialized investments are incurred, the initial winning bidder realizes no advantage over nonwinners. Although it may continue to supply for a long time, that is only because, in effect, it is continuously meeting competitive bids from qualified rivals. Rivals cannot be presumed to operate on a parity, however, once substantial investments in transaction-specific assets are put in place. Winners in such circumstances enjoy advantages over nonwinners, which is to say that parity is upset. Accordingly, what was a large numbers bidding condition at the outset is effectively transformed into one of bilateral supply thereafter. This fundamental transformation has pervasive contracting consequences.[108]
>
> The reason why significant reliance investments in durable, transaction-specific assets introduces contractual asymmetry between the winning bidder on the one hand and nonwinners on the other is that economic values would be sacrificed if the ongoing supply relation were to be terminated. Faceless contracting is thereby supplanted by contracting in which the pairwise identity of the parties matters. Occasionally the identity of the parties is important from the very outset, as when a buyer induces a supplier to invest in specialized physical capital of a transaction-specific kind. Inasmuch as the value of that capital in other uses is, by definition, much smaller than the specialized use for which it has been intended, the supplier is effectively committed to the transaction to a significant degree. The effect is often symmetrical, moreover, in that the buyer cannot turn to alternative sources of supply and obtain the item on favorable terms, since the cost of supply from unspecialized capital is presumably great.[109]

Williamson's emphasis on the limitation of the *ex post* environment of a contractual relationship underscores the enormous difficulties of negotiating a contract under conditions where asset specificity is important. Such contractual relationships may be extremely difficult to sever once made, with major financial

---

[107] *Ibid.*

[108] *Ibid.*

[109] *Ibid.*, p. 62.

consequences for both parties. It is possible, of course, that both institutional and personal trust relationships will form between the contracting parties during the execution of the contract. Personal integrity will to some extent attenuate the incentives of either party to interpret the contract language in opportunistic ways and may provide an informal mechanism for effectuating changes in the contractual relationship as contingencies emerge that were not anticipated during the *ex ante* contract negotiations. As Williamson observes, "Other things being equal, idiosyncratic exchange relations that feature personal trust will receive greater stress and will display greater adaptability."[110]

Specific and Nonspecific Costs. A standard categorization of cost in both economics and accounting is *fixed cost* and *variable cost.* In economics, fixed cost refers to the expenditures for inputs of production that do not vary with the volume of production. Such costs are sometimes called *overhead costs* or *unavoidable costs* and include such items as office rent, insurance expense, and other elements of the cost of production that are insensitive to how much is produced at any given time. By contrast, variable cost refers to expenditures on the inputs of production that vary directly with the volume of production. Such costs are also called *avoidable costs* or *direct costs* and include such expenses as the cost of labor.

In economics, the concepts of fixed and variable cost lead to the concept of a *short run total cost function* that expresses the minimized total cost of production for any level of production. In symbols, the short run total cost function may be written as

$$C(Q) = F + V(Q) \tag{B.1}$$

where C(Q) measures the minimized total cost of production expressed as a function of the level of output, Q; F measures the level of total fixed cost; and V(Q) measures total variable cost expressed as a function of the level of output produced. Given equation (B.1), the concept of the *average* (or *unit*) *cost of production* is straightforwardly defined as

$$C(Q)/Q = [F + V(Q)]/Q \tag{B.2}$$

Similarly, the important concept of the *marginal cost of production* is defined as the first derivative of equation (B.1) taken with respect to the level of output, i.e.,

$$dC(Q)/dC = dV(Q)/dC \tag{B.3}$$

Marginal cost as defined in equation (B.3) provides the cost basis for establishing economically-efficient prices for goods and services in either competitive markets or regulated industries.

---

[110] *Ibid.* pp. 62-63.

Transaction cost economics introduces a further cost distinction, namely, *specific* and *nonspecific costs*. Williamson explains the motivation for such a distinction as follows:

> It is common to distinguish between fixed and variable costs, but this is merely an accounting distinction. More relevant to the study of contracting is whether assets are redeployable or not. . . . Many assets that accountants regard as fixed are in fact redeployable, for example, centrally located general purpose buildings and equipment. Durable but mobile assets such as general purpose trucks and airplanes are likewise redeployable. Other costs that accountants treat as variable often have a large nonsalvageable part, firm-specific human capital being an illustration. . . . [111]

Thus, both total fixed cost, F, and total variable cost, V(Q), may be partitioned as

$$F \equiv F_s + F_{ns} \tag{B.4}$$

and

$$V(Q) \equiv V(Q)_s + V(Q)_{ns} \tag{B.5}$$

where the subscript s represents the portion of fixed or variable cost that is specific to a given transaction, and the subscript ns represents the remaining portion of either fixed or variable cost that is nonspecific to the transaction. Thus, two technologies of production may result in the *same* total cost of production viewed in terms of a total cost function such as equation (B.1). Nevertheless, the price of output produced by the two technologies, say, a general purpose technology and a special purpose technology, will likely differ—perhaps substantially—if the specific and nonspecific components of fixed and variable cost are substantially different.

Williamson's Hostage Concept.[112] The importance of specific and nonspecific costs is clearly illustrated in terms of Williamson's "hostage" concept. Assume that a product can be produced using either a general-purpose or a special-purpose technology. The special-purpose technology requires greater investment in transaction-specific durable assets compared to the general-purpose technology. The special-purpose technology involves both nonspecific fixed and variable costs represented by v_2, and specific fixed and variable costs represented by k. These latter costs represent the nonsalvageable (sunk) values of advance commitments in specialized assets that cannot be redeployed to alternative uses. The general-purpose technology involves only redeployable, nonspecific fixed and variable costs.

---

[111] *Ibid.*, p. 54.

[112] The following discussion is based on Oliver E. Williamson, "The Economics of Governance: Framework and Implications" in *Economics as a Process: Essays in the New Institutional Economics,* ed. Richard N. Langlois (Cambridge, U.K.: Cambridge University Press, 1986), pp. 171-202. A more detailed discussion is provided in Williamson, *The Economic Institutions of Capitalism,* Chapter 7.

In symbols, the two technologies and their costs may be written as

T_1: the general-purpose technology, with an average cost of production equal to v_1 with no specific costs; and

T_2: the special-purpose technology, with an average cost of production equal to v_2 with specific costs equal to k.

Since the special-purpose technology implies more difficult transactional problems due to asset specificity, there is an interesting choice between the two technologies only if $k + v_2 < v_1$.

If competitive market conditions prevail in the industry, then supply price using T_1 will be $p_1 = v_1$ as standard price theory would predict. Given that T_2 involves specific costs attributable to the specialized assets used in production, output price, $\bar{p}$, must necessarily exceed p_1, unless the buyer provides the seller with a safeguard to prevent the expropriation of the seller's investment in transaction-specific assets. Such a safeguard may be called a *hostage* and may consist of a performance bond posted by the buyer. If the value of the hostage offered by the buyer is just equal to the value k, then the seller may be willing to supply output to the buyer at a unit price that is just equal to v_2. This result is equivalent to the case using T_1, so long as the value of the hostage, h, is just equal to k. If, however, $h < k$, then unit price will exceed v_2 by a sufficient margin to safeguard the remainder of the seller's transaction-specific investment.[113]

Among other implications, the hostage concept provides an explanation for price discrimination, where equivalent quantities of the same product or service are priced differently for different customers. Customers that offer no hostage may be expected to pay more per unit for output that is produced using a T_2 technology compared to customers who do offer a hostage to the seller. Moreover, the same concept provides a cost-based explanation for nonlinear pricing, where customers are offered a choice of alternative two-part tariffs with differing flat fee and usage-sensitive price components reflecting different values for h. Ordinarily, a cost basis for such nonlinear rate structures is not provided in traditional statements of the theory that support such discriminatory pricing schemes.[114]

---

[113] Williamson's hostage concept is a special application of Williamson's simple contracting schema shown in Figure 3 of the text.

[114] A clear discussion of nonlinear pricing as applied in the U.S. telecommunications industry is provided by Bridger M. Mitchell and Ingo Vogelsang, *Telecommunications Pricing: Theory and Practice* (Cambridge, U.K.: Cambridge University Press, 1991), Chapter 5.